Preparing for the Big One: Prototyping Exascale Compute Clusters using Ordinary Household Objects
Company Overview
A Sample of our 2500+ Customers

National Labs

Aerospace/Defense

 Universities/Institutions

Enterprise
Complete Cluster Expertise

- Applications and Workload Knowledge
- Systems/Storage
- High-Performance Interconnects
- Scyld ClusterWare and Scyld Integrated Management Framework (IMF)
- Compilers and Tools
- World Class 24x7 Support

- Rack Integration
  > Servers, storage, power and networking
- Software Install and Configuration
  > Scyld ClusterWare, schedulers, developer tools,
  > Applications
  > Special-purpose nodes
  > Custom software installation
- Solution Testing
  > System level burn-in
  > Full cluster testing
  > Custom testing
Complete Factory Integration

- Production Lines include Assembly, Pre-test, Burn-in, Aft-test, Packing station.
- Full System Integration and Test: End-to-end testing of complete compute environment
- Integrated IT Systems for Quality and Support Configurator => CRM => SAP => SupportForce

- 3 production lines
- Fremont CA: 52,000 square feet.
- 9,000 units per month doable
- Support BTO and CTO for server and storage products
Technology Review
Background and Motivation

- High Performance Computing is a moving target
  - Users always want jobs to run faster
  - And larger jobs to be possible
- Always trying new things
  - Don Becker and Thomas Sterling invent Beowulf clusters of commodity parts
  - Infiniband high performance networking
  - NUMA memory architectures
  - Computing with GPU's
- Computing with APU's
“Everything old is new again”

- “Vector” coprocessors
  - Cray-1
  - nVidia M2090
  - AMD Firestream 9370

- “Memory” bottlenecks
  - Tape sort
  - PCI-Express
  - STREAM benchmark
What's Next?

- CPU with integrated graphics
- AMD Fusion Accelerated Processing Unit (APU)
  - Multi-core processor
  - Integrated Memory Controller
  - Integrated Graphics Controller (GPU)
  - Integrated I/O – PCI Express Root
- And ...
  - Laptop power budget
  - Laptop “Feature” list
    - x8 memory
    - No ECC
    - Soldered on board
    - Custom heat sinks
    - Laptop form factor
    - x4 PCI-Express
What Our Customers Want

- Desktop and server power budgets
  - Llano A8-3850 has 100W budget

- Standard Socket
  - FM1 (Exactly like AM2, AM3, and F only different)
  - Uses same heatsink retention

- High Performance I/O
  - Llano A8-3850 has 20 PCI-Express Gen 2 (x8,x8,x4)

- Standard Rackmount Chassis
  - Uh oh...
  - ATX, Flex-ATX, micro-ATX form factors for desktops
  - Time to get creative
Creative Use of Household Objects

- Cardboard cookie sheet
- Leftover bits from other systems
  - 80Plus power supply
  - 3-pin, 80mm fan
  - Serial cable circa 2004
  - DDR IB card
- New ingredients
  - 1866MHz memory
  - APU and motherboard
A (Nearly) Finished First Article

- 2U Chassis
  - Low Profile slots
  - 80Plus GOLD
  - Plenty of room

- Yellow buttons
Challenges

- **Management**
  - Desktop boards don't have BMC or IPMI
  - Use PDU with switched outlets
  - Use Penguin Remote Control circa 2004
    - Integrated with serial console
    - Power switch, Reset switch, Identification

- **Power draw**
  - PDU has per outlet power monitoring
  - Need fan speed control
    - ~90W @Idle w/o FSC, ~60W @Idle w/FSC
  - Detailed internal power rail monitoring
Engineering a New Product

- Low effect of OPamps
- Opt. Ethernet 10/100 RJ45
- USB header 2x5
- FTDI USB

- Voltage & Current 1% or better Accuracy/Resolution
- Voltage/CURRENT
- Opt. FAN, HDM & Capture Component (Min 3) 4H
- HDD (Min 1)
- Stacked under multi beamer

- проект 120K
- 16 Ch / MHz
- FRONT PWR: 120uW
- 33 ± 3% 100 mHz x 5K
- July 30 Aug 15
- 120 units
Challenges (cont.)

- BIOS bugs
  - MTRR issues (BIOS and kernel)
  - VESA BIOS color palette
- Hardware drivers
  - SDK for GPU not yet released
- Hardware
  - Can desktop Ethernet NIC survive cluster pounding?
- Memory performance
  - 1866MHz RAM requires “one DIMM per channel”
Next Steps

- Sample systems to customer
- Deliver 106 node cluster to customer by end of fiscal year
- Assess market potential for Altus 2A00

- Profit!
Thank you!

Contact: